UC3M Working Papers
Statistics and Econometrics
15-03
ISSN 2387-0303
March 2015

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-48

Two-sample Hotelling's T^2 statistics based on the functional Mahalanobis semi-distance

Esdras Joseph[a], Pedro Galeano[b] and Rosa E. Lillo[c]

Abstract

The comparison of the means of two independent samples is one of the most popular problems in real-world data analysis. In the multivariate context, two-sample Hotelling's T^2 frequently used to test the equality of means of two independent Gaussian random samples assuming either the same or a different covariance matrix. In this paper, we derive two-sample Hotelling's T^2 from two functional distributions. The statistics that we propose are based on the functional Mahalanobis semi-distance and, under certain conditions, their asymptotic distributions are chi-squared, regardless the distribution of the functional random samples. Additionally, we provide the link between the two-sample Hotelling's T^2 semi-distance and statistics based on the functional principal components semi-distance. A Monte Carlo study indicates that the two-sample Hotelling's T^2 of power those based on the functional principal components semi-distance. We analyze a data set of daily temperature records of 35 Canadian weather stations over a year with the goal of testing whether or not the mean temperature functions of the stations in the Eastern and Western Canada regions are equal. The results appear to indicate differences between both regions that are not found with statistics based on the functional principal components semi-distance.

Keywords: Functional Behrens-Fisher problem; Functional data analysis; Functional Mahalanobis semi-distance; Functional principal components semi-distance; Hotelling's T^2 statistics; Two-sample problems.

[a] Corresponding author. Department of Statistics, Universidad Carlos III de Madrid. E-mail: ejoseph@est-econ.uc3m.es

[b] Department of Statistics, Universidad Carlos III de Madrid. E-mail: pedro.galeano@uc3m.es

[c] Department of Statistics, Universidad Carlos III de Madrid. E-mail: rosaelvira.lillo@uc3m.es

Acknowledgements: the authors acknowledge financial support from the Spanish Ministry of Economy and Competition, research project ECO2012-38442.

Two-sample Hotelling's T^2 statistics based on the functional Mahalanobis semi-distance

Esdras Joseph, Pedro Galeano and Rosa E. Lillo
Departamento de Estadística
Universidad Carlos III de Madrid

Abstract

The comparison of the means of two independent samples is one of the most popular problems in real-world data analysis. In the multivariate context, two-sample Hotelling's T^2 statistics are frequently used to test the equality of means of two independent Gaussian random samples assuming either the same or a different covariance matrix. In this paper, we derive two-sample Hotelling's T^2 statistics for testing the equality of means in two samples independently drawn from two functional distributions. The statistics that we propose are based on the functional Mahalanobis semi-distance and, under certain conditions, their asymptotic distributions are chi-squared, regardless the distribution of the functional random samples. Additionally, we provide the link between the two-sample Hotelling's T^2 statistics based on the functional Mahalanobis semi-distance and statistics based on the functional principal components semi-distance. A Monte Carlo study indicates that the two-sample Hotelling's T^2 statistics outperform in terms of power those based on the functional principal components semi-distance. We analyze a data set of daily temperature records of 35 Canadian weather stations over a year with the goal of testing whether or not the mean temperature functions of the stations in the Eastern and Western Canada regions are equal. The results appear to indicate differences between both regions that are not found with statistics based on the functional principal components semi-distance.

Keywords: Functional Behrens-Fisher problem; Functional data analysis; Functional Mahalanobis semi-distance; Functional principal components semi-distance; Hotelling's T^2 statistics; Two-sample problems.

1 Introduction

Functional data consist of observed functions or curves evaluated at a finite interval of the real line. In a conceptual sense, functional data are intrinsically infinite dimensional and thus, classical methods designed for multivariate observations are no longer applicable. Consequently, there is a need to develop special techniques for this type of data. The books by Ramsay and Silverman (2005) and Ferraty and Vieu (2006) offer comprehensive introductions to FDA and its applications. Horváth and Kokoszka (2012) review some recent developments on inference for functional data.

In this paper, we focus on the problem of testing the equality of mean functions in two random samples independently drawn from two functional distributions. In the multivariate context, the two-sample Hotelling's T^2 statistic is frequently used to test the equality of means of two independent Gaussian random samples with the same covariance matrix which it is the multivariate analogue of the two sample t-test in the univariate case. Under the null hypothesis of equality of means, the Hotelling's T^2 statistic has a scaled F distribution. If equality of covariance matrices is not assumed, the testing issue is known as the multivariate Behrens-Fisher problem although the two-sample Hotelling's T^2 statistic is still used. In this case, several approximated scaled F distributions for the T^2 statistic under the null hypothesis have been proposed, see Rencher (1998,2000), for instance. The common point of the two statistics, that is, assuming that the covariance matrices are equal or that they are different, is that the two-sample Hotelling's T^2 statistics are just the squared Mahalanobis distance between the sample means of both random samples.

Few approaches have been proposed so far to test whether the mean functions of two functional samples are equal. For instance, Fan and Lin (1998) developed tests for comparing the means of two functional samples based on the adaptive Neyman test and wavelet thresholding techniques. Cuevas et al. (2004) proposed an ANOVA test for comparing the means of multiple samples of functional data based on the L^2-norm. Benko et al. (2009) developed bootstrap procedures for testing the equality of mean functions of two functional random samples, their functional principal components (FPCs), and their associated eigenvalues and eigenfunctions. Zhang, Peng and Zhang (2010) and Zhang, Liang and Xiao (2011) proposed a L^2-norm based statistic to test for the equality of mean functions of two Gaussian processes with possibly unequal covariance operators and derived the distributions of the proposed test statistic under the null hypothesis and a sequence of local alternatives. Finally, Horváth and Kokoszka (2012) presented procedures for testing the equality of the means in two independent functional random samples based on the functional principal components semi-distance between the sample means of the two functional samples. The asymptotic distribution of the statistic derived in this way converges, under the null hypothesis, to weighted sums of squares of independent standard Gaussians. Alternatively, to avoid the use of the weighted asymptotic distribution, Horváth and Kokoszka (2012) also proposed a normalized version of the statistic based on the functional principal components semi-distance that has a chi-square limit. These inferential procedures were extended to the case of functional time series in Horváth et al. (2013).

As mentioned previously, in the multivariate case, the two-sample Hotelling's T^2 statistic is just the squared Mahalanobis distance between the sample means of both samples. Recently, the Mahalanobis distance for multivariate observations proposed by Mahalanobis (1936) has been extended to the functional framework in Galeano et al. (2014), where it is shown as a useful tool in supervised classification problems. In this paper, we derive two-sample Hotelling's T^2 statistics based on the functional Mahalanobis semi-distance assuming either a common or a different covariance operator for the random samples following the ideas developed in the multivariate context. These statistics have asymptotically chi-squared distributions under the null hypothesis of equality of means and, contrary to the multivariate case, it is not necessary to consider the hypothesis of Gaussianity for the two populations. In particular, we show that the test statistics derived in terms of the Mahalanobis semi-distance coincide with the normalized test statistic proposed by Horváth and Kokoszka (2012) although, these authors did not consider the functional Mahalanobis semi-distance in the development of their normalized statistic. Therefore, this paper establishes the link between the Hotelling's T^2 statistic in the multivariate and functional settings.

Several Monte Carlo simulations are carried out to examine the performance of the test statistics based on the functional Mahalanobis semi-distance and the functional principal components semi-distance in scenarios previously considered in Galeano et al. (2014). The obtained results suggest that the test statistics based on the functional Mahalanobis semi-distance clearly outperform the statistics based on the functional principal components semi-distance in terms of power, at least in the considered scenarios. The obtained results appear to diverge from those of the simulation study found in Horváth and Kokoszka (2012), who indicated that neither of the two tests statistics clearly dominates the other for their simulated Gaussian data. Additionally, the analysis of a real data example from climatology suggests that the test statistic based on the functional Mahalanobis semi-distance might be more powerful than the one based on the functional principal components semi-distance.

The rest of this paper is organized as follows. Section 2 introduces some preliminaries needed to define properly the statistics associated to the homogeneity test. Section 3 introduces the statistics based on the functional Mahalanobis semi-distance for testing the equality of mean functions in two independent random samples and describes their asymptotic behavior. Sections 4 and 5 evaluate the performance of the procedures proposed in Section 3 by means of a simulation study and a real data application. Finally, some conclusions are drawn in Section 6.

2 Preliminaries

The aim of this section is to briefly review the multivariate Hotelling's T^2 statistics to motivate their extension to the functional framework. We also present some useful tools of the FDA necessary for the developments in Section 3.

2.1 Multivariate Hotelling's T^2 statistics

Let $\mathbf{x}_{11}, \ldots, \mathbf{x}_{1n_1}$ and $\mathbf{x}_{22}, \ldots, \mathbf{x}_{2n_2}$ be two random samples independently drawn from two multivariate Gaussian distributions with means $\mathbf{m}_{\mathbf{x}_1}$ and $\mathbf{m}_{\mathbf{x}_2}$ and positive definite covariance matrices $\mathbf{C}_1$ and $\mathbf{C}_2$, respectively. The aim is to test:

$$H_0 : \mathbf{m}_{\mathbf{x}_1} = \mathbf{m}_{\mathbf{x}_2} \qquad vs. \qquad H_A : \mathbf{m}_{\mathbf{x}_1} \neq \mathbf{m}_{\mathbf{x}_2}. \tag{1}$$

Let $\widehat{\mathbf{m}}_{\mathbf{x}_1} = \frac{1}{n_1}\sum_{i=1}^{n_1} \mathbf{x}_{1i}$ and $\widehat{\mathbf{m}}_{\mathbf{x}_2} = \frac{1}{n_2}\sum_{j=1}^{n_2} \mathbf{x}_{2j}$ be the sample means of the two random samples, respectively. The Multivariate Hotelling's T^2 statistic for the test (1) is given by:

$$T^2 = d_M\left(\widehat{\mathbf{m}}_{\mathbf{x}_1}, \widehat{\mathbf{m}}_{\mathbf{x}_2}\right)^2, \tag{2}$$

where $d_M\left(\widehat{\mathbf{m}}_{\mathbf{x}_1}, \widehat{\mathbf{m}}_{\mathbf{x}_2}\right)$ is the Mahalanobis distance between $\widehat{\mathbf{m}}_{\mathbf{x}_1}$ and $\widehat{\mathbf{m}}_{\mathbf{x}_2}$ defined as:

$$d_M\left(\widehat{\mathbf{m}}_{\mathbf{x}_1}, \widehat{\mathbf{m}}_{\mathbf{x}_2}\right)^2 = (\widehat{\mathbf{m}}_{\mathbf{x}_1} - \widehat{\mathbf{m}}_{\mathbf{x}_2})' \widehat{\mathbf{C}}_{12}^{-1} (\widehat{\mathbf{m}}_{\mathbf{x}_1} - \widehat{\mathbf{m}}_{\mathbf{x}_2}),$$

and $\widehat{\mathbf{C}}_{12}$ is an estimate of the covariance matrix of $\widehat{\mathbf{m}}_{\mathbf{x}_1} - \widehat{\mathbf{m}}_{\mathbf{x}_2}$ defined depending on whether $\mathbf{C}_1$ and $\mathbf{C}_2$ are assumed to be equal or not. On the one hand, if $\mathbf{C}_1 = \mathbf{C}_2 = \mathbf{C}$, the covariance matrix of $\widehat{\mathbf{m}}_{\mathbf{x}_1} - \widehat{\mathbf{m}}_{\mathbf{x}_2}$ is given by:

$$\mathbf{C}_{12} = \frac{n_1 + n_2}{n_1 n_2} \mathbf{C},$$

that can be estimated with:

$$\widehat{\mathbf{C}}_{12} = \frac{n_1 + n_2}{n_1 n_2} \widehat{\mathbf{C}}, \tag{3}$$

where $\widehat{\mathbf{C}}$ in (3) is the pooled covariance matrix given by:

$$\widehat{\mathbf{C}} = \frac{1}{n_1 + n_2 - 2} \left((n_1 - 1)\, \widehat{\mathbf{C}}_1 + (n_2 - 1)\, \widehat{\mathbf{C}}_2 \right),$$

and $\widehat{\mathbf{C}}_1$ and $\widehat{\mathbf{C}}_2$ are the sample covariance matrices of $\mathbf{C}_1$ and $\mathbf{C}_2$ based on the two random samples, respectively, given by:

$$\widehat{\mathbf{C}}_j = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} (\mathbf{x}_{ji} - \widehat{\mathbf{m}}_{\mathbf{x}_j})(\mathbf{x}_{ji} - \widehat{\mathbf{m}}_{\mathbf{x}_j})', \tag{4}$$

for $j = 1, 2$, respectively. Then, if T_C^2 denotes the multivariate Hotelling's T^2 statistic in (2) where $\widehat{\mathbf{C}}_{12}$ is given in (3), $\frac{n-p-1}{p(n-2)} T_C^2$ follows a F distribution with p and $n - p - 1$ degrees of freedom under the null hypothesis of equality of means given in (1). On the other hand, if $\mathbf{C}_1 \neq \mathbf{C}_2$, the covariance matrix of $\widehat{\mathbf{m}}_{\mathbf{x}_1} - \widehat{\mathbf{m}}_{\mathbf{x}_2}$ is given by:

$$\mathbf{C}_{12} = \frac{1}{n_1} \mathbf{C}_1 + \frac{1}{n_2} \mathbf{C}_2,$$

that can be estimated through:

$$\widehat{\mathbf{C}}_{12} = \frac{1}{n_1}\widehat{\mathbf{C}}_1 + \frac{1}{n_2}\widehat{\mathbf{C}}_2, \tag{5}$$

where $\widehat{\mathbf{C}}_1$ and $\widehat{\mathbf{C}}_2$ are defined in (4). Then, if T_D^2 denotes the multivariate Hotelling's T^2 statistic in (2) where $\widehat{\mathbf{C}}_{12}$ is given in (5), the distribution of T_D^2 under the null hypothesis in (1) has been approximated with several scaled F distributions, see James (1954), Yao (1965), Johansen (1980), Nel and van der Merwe (1986) and Kim (1992), among others.

2.2 Some notations and definitions in FDA

Let χ be a functional random variable defined in the infinite dimensional space $L^2(T)$, i.e., the space of squared integrable functions in the closed interval $T = [a, b]$. The functional variable χ has a mean function $\mu_\chi(t) = E[\chi(t)]$ and a covariance operator Γ_χ given by:

$$\Gamma_\chi(\eta) = E[(\chi - \mu_\chi) \otimes (\chi - \mu_\chi)(\eta)], \tag{6}$$

such that, for any $\eta \in L^2(T)$,

$$(\chi - \mu_\chi) \otimes (\chi - \mu_\chi)(\eta) = \langle \chi - \mu_\chi, \eta \rangle (\chi - \mu_\chi), \tag{7}$$

where $\langle \chi - \mu_\chi, \eta \rangle = \int_T (\chi(t) - \mu_\chi(t))\, \eta(t) dt$, is the usual inner product in $L^2(T)$.

If $E\left[\|\chi\|_2^2\right]$ is finite, where $\|.\|_2$ denotes the usual norm in $L^2(T)$, then Γ_χ in (6) is a compact operator, see Mas (2007), for instance. Under this assumption, there exists a sequence of non-negative eigenvalues of Γ_χ, denoted by $\lambda_1 > \lambda_2 > \cdots$, where $\sum_{k=1}^\infty \lambda_k < \infty$, and a set of orthonormal eigenfunctions of Γ_χ, denoted by $\psi_1, \psi_2, \ldots$ such that $\Gamma_\chi(\psi_k) = \lambda_k \psi_k$, for $k = 1, 2, \ldots$ The set of eigenfunctions $\psi_1, \psi_2, \ldots$ form an orthonormal basis in $L^2(T)$ and allows Γ_χ to be written as:

$$\Gamma_\chi(\eta) = \sum_{k=1}^{\infty} \lambda_k \left(\psi_k \otimes \psi_k\right)(\eta).$$

In certain circumstances, it is possible to define the inverse of the covariance operator, denoted by

Γ_χ^{-1}, as follows:

$$\Gamma_\chi^{-1}(\zeta) = \sum_{k=1}^{\infty} \frac{1}{\lambda_k}(\psi_k \otimes \psi_k)(\zeta),$$

where ζ is a function in the range of Γ_χ. However, Γ_χ^{-1} is an unbounded symmetric operator on $L^2(T)$ and extremely irregular. Hence, Mas (2007) proposed the following regularized inverse operator:

$$\Gamma_K^{-1}(\zeta) = \sum_{k=1}^{K} \frac{1}{\lambda_k}(\psi_k \otimes \psi_k)(\zeta),$$

where K is a regularization parameter. In a similar manner, we can define a regularized square root inverse operator of $\Gamma_\chi(\zeta)$ as follows:

$$\Gamma_K^{-1/2}(\zeta) = \sum_{k=1}^{K} \frac{1}{\lambda_k^{1/2}}(\psi_k \otimes \psi_k)(\zeta), \tag{8}$$

that plays a crucial role in the definition of the functional Mahalanobis semi-distance.

3 Functional two-sample Hotelling's T^2 statistics

The purpose of this section is to introduce the functional two-sample Hotelling's T^2 statistics defined through the functional Mahalanobis semi-distance proposed by Galeano et al. (2014). For that, we adapt the definitions of the two-sample Hotelling's T^2 statistics, T_C^2 and T_D^2, for multivariate data defined in Section 2.

Let χ_1 and χ_2 be two independent functional random variables defined in the infinite dimensional space $L^2(T)$, with mean functions $\mu_{\chi_1}(t) = E[\chi_1(t)]$ and $\mu_{\chi_2(t)} = E[\chi_2(t)]$ and compact covariance operators Γ_{χ_1} and Γ_{χ_2}, respectively. Therefore, χ_1 and χ_2 can be written as $\chi_1 = \mu_{\chi_1} + \epsilon_1$ and $\chi_2 = \mu_{\chi_2} + \epsilon_2$, respectively, where ϵ_1 and ϵ_2 are two independent error functional random variables defined in $L^2(T)$ with compact covariance operators Γ_{χ_1} and Γ_{χ_2}, respectively. Additionally, we assume that $E\left[\|\epsilon_j\|_2^4\right] < \infty$, for $j = 1, 2$.

Let $\chi_{11}, \ldots, \chi_{1n_1}$ and $\chi_{21}, \ldots, \chi_{2n_2}$ be two random samples independently drawn from χ_1 and χ_2, respectively. Therefore,

$$\chi_{1i}(t) = \mu_{\chi_1}(t) + \epsilon_{1i}(t), \tag{9}$$

for $1 \le i \le n_1$, and

$$\chi_{2i}(t) = \mu_{\chi_2}(t) + \epsilon_{2i}(t), \tag{10}$$

for $1 \le i \le n_2$, respectively, where $\epsilon_{11}, \dots, \epsilon_{1n_1}$ and $\epsilon_{21}, \dots, \epsilon_{2n_2}$ are two random samples independently drawn from ϵ_1 and ϵ_2, respectively. The aim is to test:

$$H_0 : \mu_{\chi_1} = \mu_{\chi_2} \qquad vs. \qquad H_A : \mu_{\chi_1} \neq \mu_{\chi_2}. \tag{11}$$

Let $\widehat{\mu}_{\chi_1} = \frac{1}{n_1}\sum_{i=1}^{n_1} \chi_{1i}$ and $\widehat{\mu}_{\chi_2} = \frac{1}{n_2}\sum_{i=1}^{n_2} \chi_{2i}$ be the sample mean functions of the two random samples, respectively, and let Γ_{12}, be the covariance operator of $\widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2}$. Similarly as in (2), we propose to test the equality of means using the functional Hotelling's T^2 statistic given by:

$$T_F^2 = d_{FM}^K(\widehat{\mu}_{\chi_1}, \widehat{\mu}_{\chi_2})^2, \tag{12}$$

where $d_{FM}^K(\widehat{\mu}_{\chi_1}, \widehat{\mu}_{\chi_2})$ is the functional Mahalanobis semi-distance between $\widehat{\mu}_{\chi_1}$ and $\widehat{\mu}_{\chi_2}$ defined as:

$$d_{FM}^K(\widehat{\mu}_{\chi_1}, \widehat{\mu}_{\chi_2})^2 = \left\langle \widehat{\Gamma}_{K,12}^{-1/2}(\widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2}), \widehat{\Gamma}_{K,12}^{-1/2}(\widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2}) \right\rangle, \tag{13}$$

where $\widehat{\Gamma}_{K,12}^{-1/2}$ is an estimate of the regularized squared root inverse covariance operator of $\widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2}$ given in (8). The estimate $\widehat{\Gamma}_{K,12}^{-1/2}$ is defined depending on whether Γ_{χ_1} and Γ_{χ_2} are assumed to be equal or not. In both cases, as shown in the Appendix, the functional Mahalanobis semi-distance in (13), and thus, the test statistic T_F^2, can be expressed as follows:

$$d_{FM}^K(\widehat{\mu}_{\chi_1}, \widehat{\mu}_{\chi_2})^2 = \sum_{k=1}^{K} \frac{\widehat{\theta}_{12k}^2}{\widehat{\lambda}_k}, \tag{14}$$

where $\widehat{\theta}_{12k} = \left\langle \widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2}, \widehat{\psi}_k \right\rangle$, for $k = 1, 2, \dots$ are the functional principal component scores with $\widehat{\psi}_1, \dots, \widehat{\psi}_K$ and $\widehat{\lambda}_1 \ge \cdots \ge \widehat{\lambda}_K$ being the eigenfunctions and associated eigenvalues, respectively, of $\widehat{\Gamma}_{12}$, an estimate of Γ_{12}, that will be given below. Consequently, the functional Hotelling's T^2 statistic T_F^2 in (12), can be written using the expression in (14) that, as mentioned before, depends on whether Γ_{χ_1} and Γ_{χ_2} are assumed to be equal or not.

On the one hand, if $\Gamma_{\chi_1} = \Gamma_{\chi_2} = \Gamma_{\chi}$, the covariance operator of $\widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2}$, is given by:

$$\Gamma_{12} = \frac{n_1 + n_2}{n_1 n_2} \Gamma_{\chi},$$

that can be estimated with:

$$\widehat{\Gamma}_{12} = \frac{n_1 + n_2}{n_1 n_2} \widehat{\Gamma}_{\chi},$$

where $\widehat{\Gamma}_{\chi}$ is the pooled covariance operator given by:

$$\widehat{\Gamma}_{\chi}(\eta) = \frac{1}{n_1 + n_2 - 2} \Big((n_1 - 1)\, \widehat{\Gamma}_{\chi_1}(\eta) + (n_2 - 1)\, \widehat{\Gamma}_{\chi_2}(\eta) \Big),$$

for $\eta \in L^2(T)$, and $\widehat{\Gamma}_{\chi_1}$ and $\widehat{\Gamma}_{\chi_2}$ being the sample covariance operators of Γ_{χ_1} and Γ_{χ_2} based on the two random samples, respectively, given by:

$$\widehat{\Gamma}_{\chi_j}(\eta) = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} \left\langle \chi_{ji} - \widehat{\mu}_{\chi_j}, \eta \right\rangle (\chi_{ji} - \widehat{\mu}_{\chi_j}), \tag{15}$$

for $j = 1, 2$, respectively. Now, eigenfunctions of $\widehat{\Gamma}_{12}$ are those of $\widehat{\Gamma}_{\chi}$, while the associated eigenvalues are $\frac{n_1+n_2}{n_1 n_2}$ times those of $\widehat{\Gamma}_{\chi}$. The statistic (14) derived in this way is the functional Hotelling's T^2 statistic assuming a common covariance operator for both samples and will be denoted by T^2_{FC}.

On the other hand, if $\Gamma_{\chi_1} \neq \Gamma_{\chi_2}$, the covariance operator of $\widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2}$ is given by:

$$\Gamma_{12} = \frac{1}{n_1} \Gamma_{\chi_1} + \frac{1}{n_2} \Gamma_{\chi_2},$$

that can be estimated through:

$$\widehat{\Gamma}_{12} = \frac{1}{n_1} \widehat{\Gamma}_{\chi_1} + \frac{1}{n_2} \widehat{\Gamma}_{\chi_2}, \tag{16}$$

where $\widehat{\Gamma}_{\chi_1}$ and $\widehat{\Gamma}_{\chi_2}$ are given in (15). Nevertheless, (16) is not the empirical covariance operator of a functional sample, as occurs in the previous case. Thus, eigenfunctions and eigenvalues of $\widehat{\Gamma}_{12}$ in (16) cannot be computed from a data set built in terms of the initial data set. For that reason, we will use the following bootstrap procedure to estimate eigenfunctions and eigenvalues of $\widehat{\Gamma}_{12}$:

Step 1 Let $b = 1$.

Step 2 Obtain a random sample with replacement from $\chi_{11}, \ldots, \chi_{1n_1}$ and another one from $\chi_{21}, \ldots, \chi_{2n_2}$. Denote both bootstrap samples by $\chi_{11}^b, \ldots, \chi_{1n_1}^b$ and $\chi_{21}^b, \ldots, \chi_{2n_2}^b$, respectively.

Step 3 Obtain the functional sample means of the bootstrap samples, denoted by $\widehat{\mu}_{\chi_1}^b$ and $\widehat{\mu}_{\chi_2}^b$, respectively and their difference $\widehat{\mu}_{12}^b = \widehat{\mu}_{\chi_1}^b - \widehat{\mu}_{\chi_2}^b$.

Step 4 Repeat Steps 2 and 3 B times to obtain B bootstrap samples $\widehat{\mu}_{12}^b$, for $b = 1, \ldots, B$. Then, the covariance operator Γ_{12} is estimated with the sample covariance operator of $\widehat{\mu}_{12}^1, \ldots, \widehat{\mu}_{12}^B$, from which we obtain the set of estimated eigenfunctions and associated eigenvalues needed to compute (14).

The statistic (14) derived in this way is the functional Hotelling's T^2 statistic assuming different covariance operators for both samples and will be denoted by T_{FD}^2.

To analyze the convergence of the statistics T_{FC}^2 and T_{FD}^2 under the null and alternative hypotheses, we briefly review the statistics given in Horváth and Kokoszka (2012) to test (11). Firstly, Horváth and Kokoszka (2012) proposed to use the statistic based on the L^2 distance defined as:

$$U = \frac{n_1 n_2}{n_1 + n_2} d_2(\widehat{\mu}_{\chi_1}, \widehat{\mu}_{\chi_2})^2 = \frac{n_1 n_2}{n_1 + n_2} \int_T (\widehat{\mu}_{\chi_1}(t) - \widehat{\mu}_{\chi_2}(t))^2 \, dt. \tag{17}$$

Under the conditions given at the beginning of this section and assuming that

$$\frac{n_1}{n_1 + n_2} \to \nu$$

with some $0 \leq \nu \leq 1$, the asymptotic distribution of (17) under the null hypothesis is the distribution of $\sum_{k=1}^{\infty} \tau_k z_k^2$, where $\tau_1 \geq \tau_2 \geq \ldots$ denotes the eigenvalues of the operator $(1 - \nu)\,\Gamma_{\chi_1} + \nu \Gamma_{\chi_2}$ and z_k are independent standard Gaussian random variables. As these eigenvalues are unknown, alternatively, Horváth and Kokoszka (2012) considered the statistic:

$$U_F = \frac{n_1 n_2}{n_1 + n_2} d_{PC}^K(\widehat{\mu}_{\chi_1}, \widehat{\mu}_{\chi_2})^2 = \sum_{k=1}^{K} \widehat{\theta}_{12k}^2, \tag{18}$$

where d_{PC}^K denoted the functional principal components semi-distance introduced in Ferraty and Vieu (2006) and $\widehat{\theta}_{121} \geq \widehat{\theta}_{122} \geq \ldots$ are the functional principal component scores of $\widehat{\Gamma}_{12}$. In other words, the idea is to replace in (17) the L^2 distance with the functional principal components semi-distance. The asymptotic distribution of the statistic U_F in (18) under the null hypothesis is the distribution of $\sum_{k=1}^{K} \tau_k z_k^2$, for which critical values can be obtained by simulation. Nevertheless, to avoid the use of simulation, Horváth and Kokoszka (2012) proposed a normalized version of U_F given by:

$$NU_F = \sum_{k=1}^{K} \frac{\widehat{\theta}_{12k}^2}{\widehat{\lambda}_k},$$

that has an asymptotic χ_K^2 distribution, see Theorem 5.3 in Horváth and Kokoszka (2012). Now, the statistic NU_F is just the functional Hotelling's T^2 statistic in (14) that, consequently, inherits the χ_K^2 asymptotic distribution. Additionally, Theorem 5.4 in Horváth and Kokoszka (2012) establishes the consistency of the NU_F statistic to reject the null hypothesis if the means are different. For that, it is necessary to assume that $\mu_{\chi_1} - \mu_{\chi_2}$ is not orthogonal to the linear span of $\psi_1, \ldots, \psi_K$. This consistency result is also inherited by the functional Hotelling's T^2 statistics. In the following, we denote by U_{FC} and U_{FD}, the statistic U_F when assuming a common or a different covariance operator of the random samples under analysis.

Finally, the threshold parameter K deserves some comments. In practice, the functional Hotelling's T^2 statistics T_{FC}^2 and T_{FD}^2, as well as the statistics U_{FC} and U_{FD} based on the functional principal components semi-distance, can be used to solve the testing problem with several values of K. Then, one can compare the results of the tests. However, it would be advisable to define a procedure that chooses an appropriate value of K to make a unique decision when this hypothesis test is applied to real data. Galeano et al. (2014) propose to select K to compute the functional Mahalanobis semi-distance in classification problems by cross-validation. However, this method can not be easily extended in the hypothesis testing framework. Alternatively, we choose the threshold value K via the cumulative percentage of total variance (CPV), that is the classical approach for determining the number of sample principal components to retain. The cumulative percentage of total variance is defined as follows:

$$CPV\left(k\right) = \frac{\sum_{j=1}^{k} \widehat{\lambda}_j}{\sum_{j=1}^{k_{\max}} \widehat{\lambda}_j}, \tag{19}$$

where $\widehat{\lambda}_j$ are the eigenvalues of $\widehat{\Gamma}_{12}$ and $k_{\max}$ is the total number of estimated eigenvalues. The CPV in (19) is an increasing function that tends to 1. Then, we select the value of K as the value of k from which the function CPV grows very slowly to 1. This is the method that we use in the simulated and real data examples in Sections 4 and 5.

4 Empirical Results

This section illustrates the performance of the test statistics presented in Section 3 through several Monte Carlo simulations. In particular, we compare the empirical sizes and powers of the test statistics based on the functional Mahalanobis semi-distance, T^2_{FC} and T^2_{FD}, with those of the test statistics based on the functional principal components semi-distance, U_{FC} and U_{FD}, when the covariance operators of the two random samples are assumed to be equal and when this is not assumed. In practice, the curves are usually observed with noise and in a finite set of sampling points that could be unequally spaced and different among the sample units. Therefore, the first step is to convert raw discrete data points into smooth functions.

4.1 Smoothing with basis functions

Usually, a functional dataset has the form $\{\chi_i^*(t_{i,q}) : i = 1, \ldots, n \text{ and } q = 1, \ldots, Q_i\}$ where n is the number of observed curves and Q_i is the number of observations of the noisy curve χ_i^* at points $t_{i,1}, \ldots, t_{i,Q_i}$. Thus, the fisrt step in FDA is to reconstruct the functional form of the sample curves from their discrete observations. One of the usual approaches to solve this problem, and the one taken in this paper, is to use basis functions. In general, a basis is a system of functions, denoted by φ_m, $m = 1, 2, \ldots$, orthogonal or not, such that $\chi_i^*(t)$, for $i = 1, \ldots, n$, can be fairly well approximated with:

$$\chi_i(t) = \sum_{m=1}^{M} \beta_{im}\varphi_m(t),$$

where β_{im}, $m = 1, \ldots, M$, are the coefficients of the expansion. The choice of the basis and the number M of basis functions to provide a smooth approximation to the observed discretized points is very important and must be done according to the characteristics of the data. The basis systems typically used are Fourier basis, for periodic data sets, and B-spline basis, for nonperiodic data sets. The simplest method to effectively estimate the coefficients of the expansion is carried out by minimizing:

$$\left(\sum_{q=1}^{Q_i} \left[\chi_i^* (t_{i,q}) - \sum_{m=1}^{M} \beta_{im} \varphi_m (t_{i,q}) \right]^2 \right)^{1/2} .$$

Once the observed data set $\{\chi_i^* (t_{i,j}) : i = 1, \ldots, n \text{ and } q = 1, \ldots, Q_i\}$ has been smoothed, we work with the smoothed functional sample $\{\chi_i (t) : i = 1, \ldots, n\}$.

We use the methods described in Ramsay and Silverman (2005) and implemented in the **R** package *fda*, see Ramsay et al. (2009), to carry out all the computations. In particular, the computation of eigenvalues and eigenfunctions of the covariance operators and functional principal component scores needed to compute the test statistics T^2_{FC}, T^2_{FD}, U_{FC} and U_{FD} are described in Section 8.4 of Ramsay and Silverman (2005).

4.2 Monte Carlo Study

In this Monte Carlo study, we generate functional data sets following the structure described in (9) and (10). In particular, we consider the functional means $\mu_{\chi_1}(t) = 20t^\rho(1 - t)$ and $\mu_{\chi_2}(t) = 20t(1 - t)^\rho$, respectively, where $\rho = 1, 1.01, 1.02, 1.03, 1.04, 1.05$. Thus, when $\rho = 1$, the null hypothesis holds, which allows us to calculate empirical sizes associated with the test statistics. However, when $\rho \neq 1$, the alternative hypothesis holds allowing the calculation of the corresponding empirical powers. Note also that the larger ρ, the more different are μ_{χ_1} and μ_{χ_2}, as plotted in Figure 1. Then, the power is a function of the parameter ρ.

First, we compare the empirical sizes and powers of the testing procedures when the covariance operators of the two random samples are equal. For that, we consider two different scenarios for the error terms. In the first scenario, we have:

$$\epsilon_1(t) = \sum_{k=1}^{\infty} \lambda_k^{1/2} z_{1k} \psi_k(t) \qquad \text{and} \qquad \epsilon_2(t) = \sum_{k=1}^{\infty} \lambda_k^{1/2} z_{2k} \psi_k(t),$$

where $\psi_k(t) = \sqrt{2} \sin((k - 0.5)\pi t)$, $t \in [0, 1]$, for $k = 1, 2, \ldots$ are the eigenfunctions of the covariance operator of the error functions with associated eigenvalues $\lambda_k = 1/(\pi(k - 0.5))^2$, for $k = 1, 2, \ldots$, and z_{1k} and z_{2k} are independent standard Gaussian distributed, for $k = 1, 2, \ldots$ Thus, ϵ_1 and ϵ_2 are two Brownian motions with a common covariance operator. In the second scenario, z_{1k}

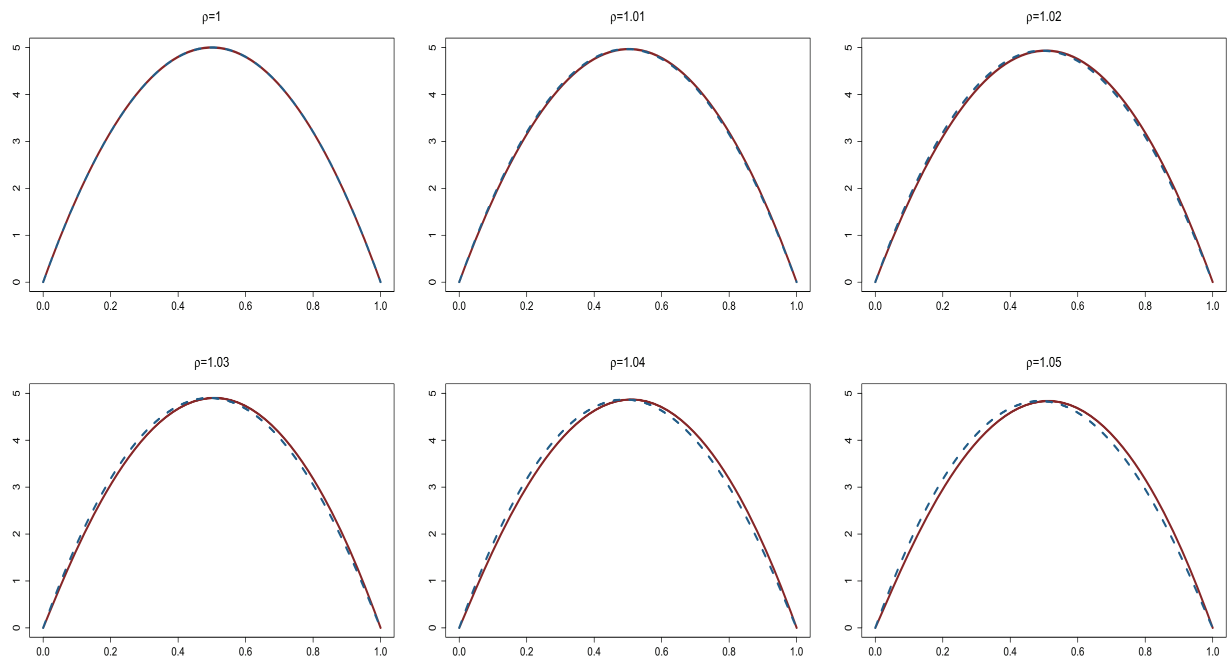


Figure 1: Mean functions for different values of ρ. In solid, first sample, and in dashed, second sample.

and z_{2k} are replaced with e_{1k} and e_{2k}, that are independent standardized exponential distributed with rate 1. We consider four configurations of sample sizes (n_1, n_2) given by $(50, 50)$, $(50, 100)$, $(100, 100)$ and $(100, 200)$, respectively. We choose these pairs in order to see how the sample sizes influence the test results.

Subsequently, 1000 data sets are generated of each scenario and pair of sample sizes. The generated functions are observed at $Q = 100$ equidistant points in the closed interval $I = [0, 1]$. Gaussian errors with mean 0 and variance 0.01 are added to each generated point. To compute the test statistics, the discrete trajectories are converted to functional observations using a B-spline basis of order 6 with 20 basis functions which seem enough to fit the data well. Figure 2 shows a data set generated from the first scenario with $\rho = 1.05$ and sample size pair $(10, 10)$ with the corresponding sample means. Note that it would be difficult to affirm through visual evaluation that the mean generating functions are different.

As mentioned in Section 3, the functional Hotelling's T^2 statistics can be computed for several values of K, for $K = 1, 2, \ldots$. Nevertheless, it would be preferable to select an appropriate threshold value K and this is done using the CPV criterion in (19). In this case, we know the true eigenvalues

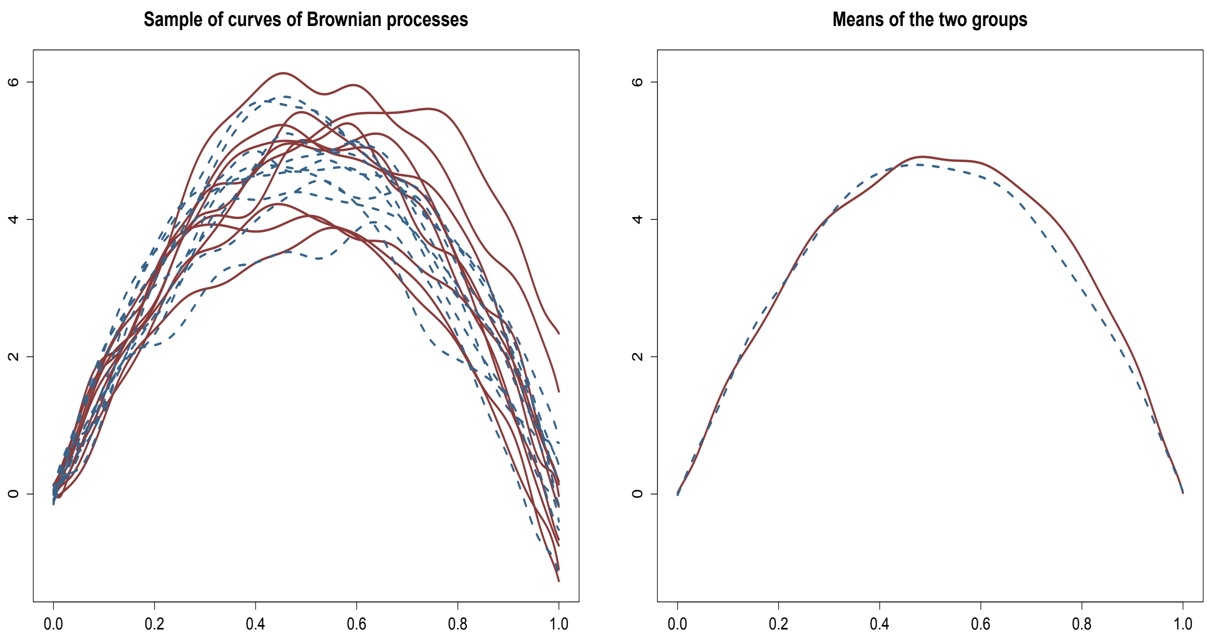


Figure 2: Left: Set of 10 functions of the Brownian Motion plus mean $\mu_{\chi_1}(t) = 20t^{1.05}(1-t)$ (solid) and another set of 10 functions of the Brownian Motion plus mean $\mu_{\chi_2}(t) = 20t(1-t)^{1.05}$ (dashed). Right: the sample functional means for the first (solid) and second (dashed) set of curves.

of the covariance operators considered in the Monte Carlo study. These eigenvalues are proportional to those of the covariance operator of the difference of the sample means so that we can use the cumulative percentages obtained from them to select an appropriate threshold K. The first ten cumulative percentages are given by 0.8216, 0.9129, 0.9458, 0.9625, 0.9717, 0.9795, 0.9843, 0.9880, 0.9908 and 0.9931, respectively. As can be seen, the cumulative percentages grow very slowly from the fifth eigenvalue. Thus, we take $CPV = 0.97$ and select the K such that the principal components explain at least the 97% of the variance. After that, we compute the T^2_{FC} and U_C statistics. Obviously, in practice (with real data) the eigenvalues of the covariance operator of the difference of sample means are unknown. Nevertheless, to take an appropriate CPV, we can use those eigenvalues estimated from the samples using the methods previously described.

The results are summarized in Figure 3 and Tables 1 and 2. On the one hand, Figure 3 shows a barplot of the values of K selected by CPV for the 1000 data sets generated for the case of $\rho = 0$ in scenario 1. As can be seen, K only takes values ranging from 4 to 7, being $K = 5$ and $K = 6$ the most frequent values. On the other hand, Tables 1 and 2 show the empirical sizes and powers

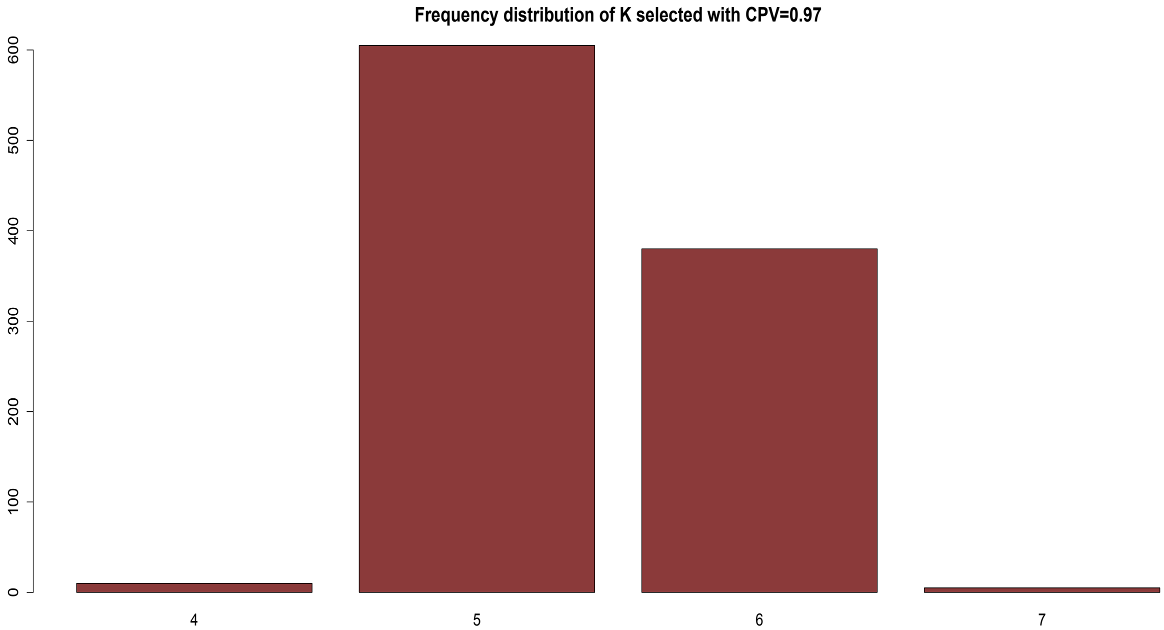


Figure 3: Values of K selected in the 1000 simulations with Gaussian processes.

of the test statistics for the two scenarios. Each cell in the tables displays the empirical size or power over the 1000 generated data sets. Empirical sizes and powers are calculated at the nominal sizes $\alpha = 0.1, 0.05, 0.01$. In view of these tables, several comments are in order. First, the empirical sizes of the two test statistics are very close to the corresponding nominal sizes in most of the cases. Indeed, the empirical sizes appears to tend to the nominal sizes as the sample sizes increase. Second, if one of the sample sizes is 50, the test statistics have empirical sizes slightly larger than the nominal sizes. Third, in terms of power, the functional Hotelling's T^2 statistic, T^2_{FC}, clearly dominates the test statistic based on the functional principal components semi-distance, U_C, in all the cases. Fourth, the functional Hotelling's T^2 test statistic has good and similar power for both Gaussian and exponential data sets suggesting that non-Gaussianity is not a drawback for T^2_{FC}. Fifth, when the parameter ρ increases, the power of U_C increases slower than that for T^2_{FC}. Sixth, the larger the sample size, the larger the power of both statistics. In summary, we conclude that the functional Hotelling's T^2 statistic appears to outperform the test statistic based on the functional principal components semi-distance in terms of power.

Next, we compare the empirical sizes and powers of T^2_{FD} and U_{FD} when the covariance operators

Table 1: Empirical sizes and powers of the functional Hotelling's T^2 statistic and the test statistic based on the functional principal components semi-distance when $\Gamma_{\chi_1} = \Gamma_{\chi_2}$ for the first scenario.

n_1	n_2	ρ	T^2_{FC} 10%	U_{FC} 10%	T^2_{FC} 5%	U_{FC} 5%	T^2_{FC} 1%	U_{FC} 1%
		1.00	0.149	0.112	0.075	0.063	0.018	0.009
		1.01	0.186	0.111	0.109	0.058	0.031	0.015
50	50	1.02	0.351	0.125	0.232	0.071	0.099	0.015
		1.03	0.660	0.179	0.537	0.099	0.316	0.023
		1.04	0.865	0.230	0.802	0.124	0.595	0.035
		1.05	0.973	0.293	0.947	0.154	0.844	0.041
n_1	n_2	ρ	T^2_{FC} 10%	U_{FC} 10%	T^2_{FC} 5%	U_{FC} 5%	T^2_{FC} 1%	U_{FC} 1%
		1.00	0.129	0.112	0.068	0.056	0.023	0.013
		1.01	0.200	0.105	0.133	0.063	0.036	0.010
50	100	1.02	0.497	0.153	0.378	0.091	0.183	0.019
		1.03	0.783	0.183	0.673	0.097	0.432	0.017
		1.04	0.951	0.283	0.898	0.141	0.757	0.039
		1.05	0.993	0.460	0.980	0.237	0.943	0.058
n_1	n_2	ρ	T^2_{FC} 10%	U_{FC} 10%	T^2_{FC} 5%	U_{FC} 5%	T^2_{FC} 1%	U_{FC} 1%
		1.00	0.109	0.101	0.054	0.048	0.013	0.007
		1.01	0.217	0.113	0.145	0.049	0.043	0.008
100	100	1.02	0.616	0.156	0.486	0.079	0.266	0.018
		1.03	0.925	0.235	0.872	0.114	0.694	0.036
		1.04	0.995	0.444	0.986	0.248	0.944	0.053
		1.05	1.000	0.678	1.000	0.410	0.998	0.119
n_1	n_2	ρ	T^2_{FC} 10%	U_{FC} 10%	T^2_{FC} 5%	U_{FC} 5%	T^2_{FC} 1%	U_{FC} 1%
		1.00	0.107	0.101	0.050	0.059	0.005	0.013
		1.01	0.263	0.114	0.172	0.056	0.062	0.012
100	200	1.02	0.709	0.182	0.602	0.101	0.360	0.018
		1.03	0.972	0.292	0.934	0.154	0.842	0.038
		1.04	0.999	0.596	0.997	0.320	0.988	0.079
		1.05	1.000	0.878	1.000	0.616	1.000	0.165

of the two random samples are different. As before, we consider two different scenarios for the error terms. In the first scenario, we have:

$$\epsilon_1(t) = \sum_{k=1}^{\infty} \lambda_{1k}^{1/2} z_{1k} \psi_k(t) \qquad \text{and} \qquad \epsilon_2(t) = \sum_{k=1}^{\infty} \lambda_{2k}^{1/2} z_{2k} \psi_k(t),$$

where $\psi_k(t) = \sqrt{2}\sin((k-0.5)\pi t)$, $t \in [0,1]$, for $k = 1, 2, \ldots$ are the eigenfunctions of the covariance operator of the error functions with associated eigenvalues $\lambda_{1k} = 1/(\pi(k-0.5))^2$ and $\lambda_{2k} = 2/(\pi(k-0.5))^2$, for $k = 1, 2, \ldots$, for the first and second random samples, respectively, and z_{1k} and z_{2k} are independent standard Gaussian distributed, for $k = 1, 2, \ldots$ Thus, ϵ_1 and ϵ_2 are

Table 2: Empirical sizes and powers of the functional Hotelling's T^2 statistic and the test statistic based on the functional principal components semi-distance when $\Gamma_{\chi_1} = \Gamma_{\chi_2}$ for the second scenario.

n_1	n_2	ρ	T^2_{FC} 10%	U_{FC} 10%	T^2_{FC} 5%	U_{FC} 5%	T^2_{FC} 1%	U_{FC} 1%
		1.00	0.113	0.111	0.065	0.053	0.017	0.019
		1.01	0.193	0.128	0.107	0.066	0.033	0.018
50	50	1.02	0.387	0.139	0.283	0.067	0.113	0.012
		1.03	0.655	0.185	0.533	0.093	0.317	0.016
		1.04	0.868	0.265	0.796	0.135	0.620	0.039
		1.05	0.970	0.347	0.953	0.175	0.850	0.033
n_1	n_2	ρ	T^2_{FC} 10%	U_{FC} 10%	T^2_{FC} 5%	U_{FC} 5%	T^2_{FC} 1%	U_{FC} 1%
		1.00	0.102	0.111	0.054	0.046	0.014	0.014
		1.01	0.205	0.108	0.127	0.059	0.035	0.013
50	100	1.02	0.480	0.115	0.332	0.056	0.146	0.007
		1.03	0.781	0.193	0.666	0.096	0.447	0.032
		1.04	0.946	0.293	0.900	0.169	0.766	0.043
		1.05	0.995	0.444	0.988	0.237	0.940	0.064
n_1	n_2	ρ	T^2_{FC} 10%	U_{FC} 10%	T^2_{FC} 5%	U_{FC} 5%	T^2_{FC} 1%	U_{FC} 1%
		1.00	0.098	0.096	0.052	0.051	0.014	0.007
		1.01	0.225	0.118	0.131	0.062	0.047	0.014
100	100	1.02	0.638	0.165	0.514	0.076	0.272	0.022
		1.03	0.913	0.266	0.851	0.137	0.685	0.038
		1.04	0.991	0.459	0.982	0.242	0.945	0.067
		1.05	1.000	0.690	1.000	0.406	0.995	0.123
n_1	n_2	ρ	T^2_{FC} 10%	U_{FC} 10%	T^2_{FC} 5%	U_{FC} 5%	T^2_{FC} 1%	U_{FC} 1%
		1.00	0.100	0.112	0.043	0.051	0.010	0.013
		1.01	0.278	0.116	0.175	0.074	0.078	0.016
100	200	1.02	0.710	0.162	0.590	0.080	0.351	0.017
		1.03	0.957	0.317	0.933	0.180	0.835	0.037
		1.04	1.000	0.566	0.999	0.309	0.989	0.088
		1.05	1.000	0.874	1.000	0.634	1.000	0.188

two Brownian motions with the same eigenfunctions but with the eigenvalues corresponding to the second error process twice those corresponding to the first error process. In the second scenario, and similarly to the case of common covariance operators, z_{1k} and z_{2k} are replaced with e_{1k} and e_{2k}, that are independent standardized exponential distributed with rate 1.

Then, 1000 data sets are generated of each pair of sample sizes and scenario with the same configurations of samples sizes and generation mechanism as in the first set of simulations. For each generated data set, we obtain $B = 1000$ bootstrap samples as explained in Section 3 allowing us to obtain the eigenfunctions and eigenvalues of the estimated covariance operator of the difference of the sample means of the two random samples. Then, in order to fix the value of CPV used in the simulation study to compute T_{FD}^2 and U_{FD}, we compute the mean bootstrap eigenvalues based on the 1000 data sets. A visual inspection of these eigenvalues for each pair of sample sizes and scenario leaded us to select $CPV = 0.97$ in all the situations. Subsequently, once the value of K has been fixed, we compute the two statistics for each generated data set.

The results are summarized in Tables 3 and 4 that show the empirical sizes and powers of the test statistics for the two scenarios. As in the previous case, each cell in the tables displays the empirical size or power calculated at the nominal sizes $\alpha = 0.1, 0.05, 0.01$ over the 1000 generated data sets. The results in terms of sizes and powers of the simulation study when the covariance operators of the random samples are different are very similar to those when the covariance operators of the random samples are the same. In particular, we would like to note that the bootstrap procedure does not appear to have a significant effect on the limit behavior of the test statistics. Finally, we repeated the study with $B = 10000$ bootstrap replications obtaining similar results, which for brevity are omitted in this paper.

5 Real data study

In this section, we compare the results obtained by the functional Hotelling's T^2 statistics and the test statistics based on the functional principal components semi-distance with the Canadian Temperature data set previously analyzed by Ramsay and Silverman (2005) and Zhang and Chen (2007), among others. The data set contains the daily temperature records of 35 Canadian weather stations over a year (365 days). As in Zhang and Chen (2007), the 35 stations have been split in three regions, resulting in 15 stations in the Eastern region, another 15 stations in the Western region and the remaining 5 stations in the Northern region. See Table 5 to see the stations assigned in each of the three regions. Following Ramsay and Silverman (2005) and Ramsay et al. (2009), the discrete observations are converted to functional observations using a Fourier series basis with 65 basis functions. Figure 4 shows the smoothed temperature curves of the Eastern (solid), Western (dashed) and Northern (dotted) weather stations and the estimated mean temperature functions of these regions. As can be seen, the mean temperature functions of the stations in the Eastern and Western regions look like similar and far from the mean temperature function of the Northern

Table 3: Empirical sizes and powers of the functional Hotelling's T^2 statistic and the test statistic based on the functional principal components semi-distance when $\Gamma_{\chi_1} \neq \Gamma_{\chi_2}$ for the first scenario.

n_1	n_2	ρ	T^2_{FD} 10%	U_{FD} 10%	T^2_{FD} 5%	U_{FD} 5%	T^2_{FD} 1%	U_{FD} 1%
		1.00	0.113	0.102	0.066	0.046	0.017	0.012
		1.01	0.156	0.103	0.089	0.058	0.024	0.013
50	50	1.02	0.304	0.118	0.20	0.064	0.084	0.015
		1.03	0.473	0.150	0.358	0.070	0.159	0.019
		1.04	0.713	0.164	0.604	0.094	0.373	0.021
		1.05	0.868	0.246	0.790	0.125	0.598	0.033
n_1	n_2	ρ	T^2_{FD} 10%	U_{FD} 10%	T^2_{FD} 5%	U_{FD} 5%	T^2_{FD} 1%	U_{FD} 1%
		1.00	0.115	0.122	0.065	0.057	0.019	0.009
		1.01	0.171	0.105	0.083	0.051	0.029	0.009
50	100	1.02	0.366	0.115	0.260	0.059	0.122	0.010
		1.03	0.654	0.169	0.531	0.087	0.308	0.020
		1.04	0.894	0.259	0.820	0.142	0.608	0.033
		1.05	0.977	0.296	0.945	0.155	0.855	0.036
n_1	n_2	ρ	T^2_{FD} 10%	U_{FD} 10%	T^2_{FD} 5%	U_{FD} 5%	T^2_{FD} 1%	U_{FD} 1%
		1.00	0.117	0.091	0.062	0.052	0.009	0.014
		1.01	0.189	0.099	0.119	0.048	0.039	0.009
100	100	1.02	0.465	0.155	0.358	0.078	0.150	0.016
		1.03	0.753	0.203	0.642	0.102	0.426	0.015
		1.04	0.934	0.289	0.891	0.143	0.761	0.034
		1.05	0.997	0.451	0.992	0.252	0.958	0.057
n_1	n_2	ρ	T^2_{FD} 10%	U_{FD} 10%	T^2_{FD} 5%	U_{FD} 5%	T^2_{FD} 1%	U_{FD} 1%
		1.00	0.100	0.110	0.052	0.058	0.017	0.013
		1.01	0.238	0.089	0.146	0.042	0.043	0.007
100	200	1.02	0.611	0.160	0.494	0.073	0.274	0.018
		1.03	0.903	0.231	0.853	0.129	0.668	0.035
		1.04	0.991	0.413	0.982	0.223	0.933	0.047
		1.05	1.000	0.688	1.000	0.416	0.994	0.121

weather stations.

Based on the reconstructed temperature curves, the objective is to test if the mean temperature functions of the Eastern and Western weather stations during the whole year are the same. We are also interested in testing if the weather stations in the Eastern and Northern and the Western and Northern regions have, respectively, the same mean temperature functions. Before performing the tests, a task that we have to carry out is to verify whether the covariance operators of the groups can be assumed to be the same, in order to choose the appropriate test statistics. For that, Figures 5 and 6 show the estimated standard deviations and covariance operators surfaces for the curves

Table 4: Empirical sizes and powers of the functional Hotelling's T^2 statistic and the test statistic based on the functional principal components semi-distance when $\Gamma_{\chi_1} \neq \Gamma_{\chi_2}$ for the second scenario.

n_1	n_2	ρ	T^2_{FD} 10%	U_{FD} 10%	T^2_{FD} 5%	U_{FD} 5%	T^2_{FD} 1%	U_{FD} 1%
		1.00	0.130	0.120	0.065	0.071	0.017	0.013
		1.01	0.156	0.129	0.092	0.066	0.021	0.019
50	50	1.02	0.273	0.116	0.174	0.069	0.056	0.023
		1.03	0.458	0.152	0.319	0.081	0.151	0.021
		1.04	0.720	0.192	0.596	0.102	0.347	0.032
		1.05	0.895	0.258	0.827	0.143	0.632	0.047
n_1	n_2	ρ	T^2_{FD} 10%	U_{FD} 10%	T^2_{FD} 5%	U_{FD} 5%	T^2_{FD} 1%	U_{FD} 1%
		1.00	0.118	0.109	0.062	0.053	0.015	0.012
		1.01	0.199	0.125	0.123	0.063	0.039	0.014
50	100	1.02	0.362	0.138	0.252	0.077	0.113	0.017
		1.03	0.659	0.181	0.523	0.096	0.330	0.021
		1.04	0.866	0.228	0.786	0.125	0.612	0.030
		1.05	0.969	0.345	0.943	0.180	0.831	0.049
n_1	n_2	ρ	T^2_{FD} 10%	U_{FD} 10%	T^2_{FD} 5%	U_{FD} 5%	T^2_{FD} 1%	U_{FD} 1%
		1.00	0.110	0.108	0.062	0.053	0.017	0.010
		1.01	0.159	0.110	0.085	0.063	0.020	0.009
100	100	1.02	0.445	0.158	0.309	0.092	0.143	0.024
		1.03	0.741	0.192	0.638	0.108	0.427	0.019
		1.04	0.943	0.299	0.910	0.169	0.777	0.048
		1.05	0.996	0.476	0.991	0.258	0.955	0.087
n_1	n_2	ρ	T^2_{FD} 10%	U_{FD} 10%	T^2_{FD} 5%	U_{FD} 5%	T^2_{FD} 1%	U_{FD} 1%
		1.00	0.115	0.108	0.060	0.054	0.012	0.020
		1.01	0.245	0.114	0.165	0.057	0.057	0.015
100	200	1.02	0.620	0.173	0.507	0.084	0.286	0.018
		1.03	0.917	0.250	0.862	0.123	0.681	0.032
		1.04	0.991	0.426	0.975	0.225	0.930	0.053
		1.05	1.000	0.710	1.000	0.447	0.995	0.134

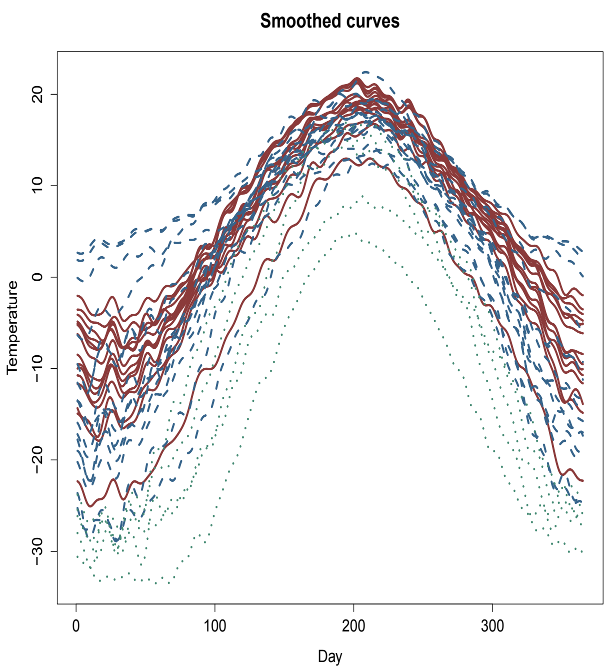
in the Eastern, Western and Northern regions, respectively, while Figure 7 show the corresponding contour plots of the estimated covariance operators. The figures show different shapes and scales suggesting that the covariance operators of the groups are different. Additionally, Figure 8 displays the eigenvalues of each estimated covariance operator that appears to move in quite different scales again leading to similar conclusions. Hence, we use the test statistics when the covariance operators of the random samples are assumed to be different.

Next, we compute the statistics T^2_{FD} and U_{FD} for $K = 1, \ldots, 15$ for the three pairs of regions with 1000 bootstrap replications. Table 6 displays the p-values of the two test statistics. As can

Table 5: Classification of the Canadian weather stations.

Eastern	St. Johns	Halifax	Sydney	Yarmouth	Charlottesville
	Fredericton	Scheffervll	Arvida	Bagottville	Quebec
	Sherbrooke	Montreal	Ottawa	Toronto	London
Western	Thunderbay	Winnipeg	The Pas	Churchill	Regina
	Pr. Albert	Uranium City	Edmonton	Calgary	Kamloops
	Vancouver	Victoria	Pr. George	Pr. Rupert	Whitehorse
Northern	Dawson	Yellowknife	Iqaluit	Inuvik	Resolute

Figure 4: Left: Daily temperature of Canada (Eastern weather stations in solid lines, Western weather stations in dashed lines and Northern weather stations in dotted lines). Right: Estimated mean temperature functions of the Eastern, Western and Northern weather stations.



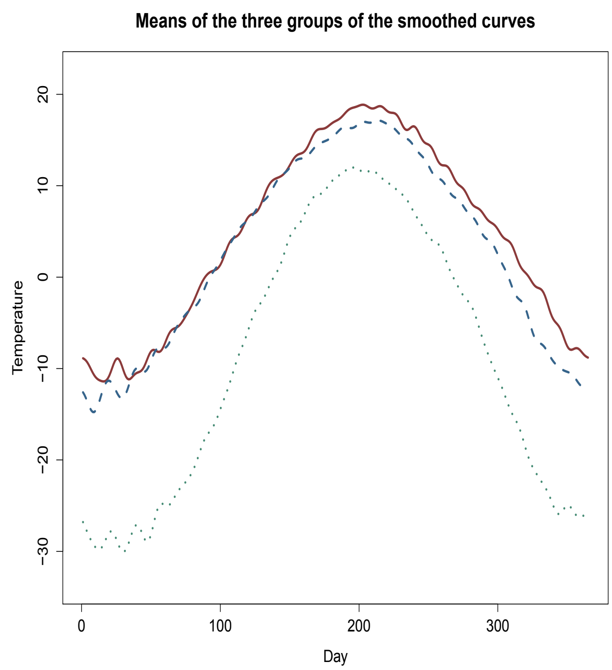
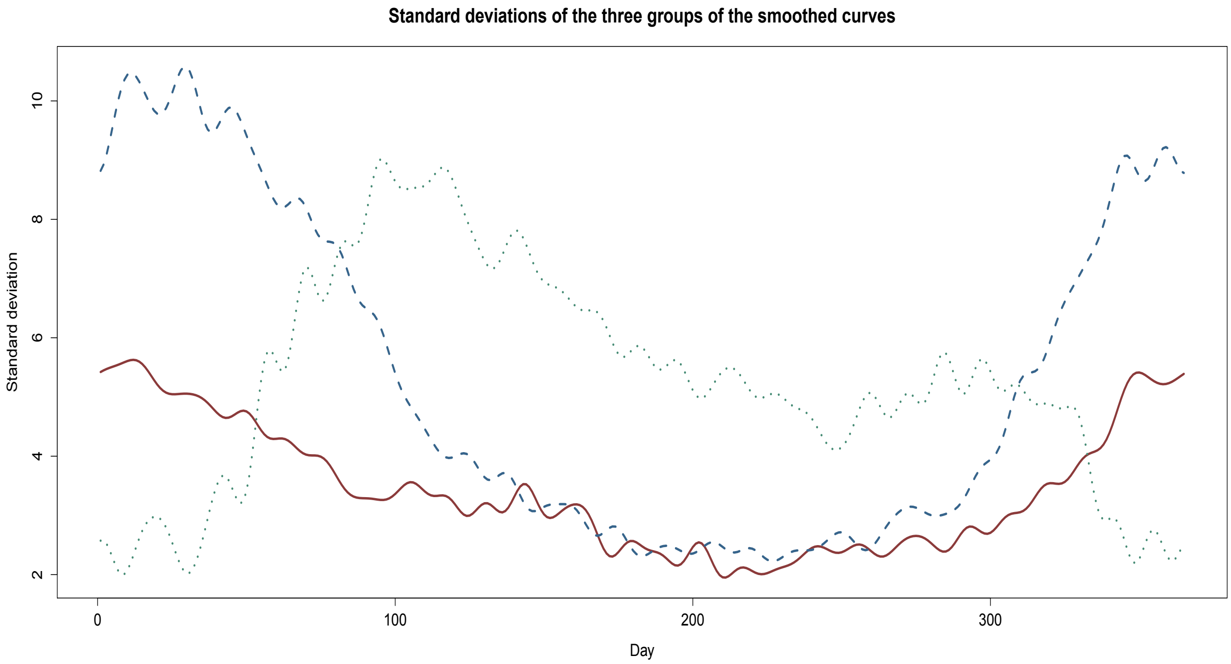


Figure 5: Estimated standard deviations of the three groups of the smoothed curves (Eastern weather stations in solid lines, Western weather stations in dashed lines and Northern weather stations in dotted lines).



be seen, both testing procedures lead to essentially the same conclusions, rejecting the equality of mean temperature functions between the Eastern and Northern regions and Western and Northern regions. However, for Eastern-Western regions, U_{FD} do not reject the null hypothesis of equality of mean functions, while T^2_{FD} reject this null hypothesis when $K > 2$. Then, we select an appropriate value of K using the cumulative percentage of total variance. For that, for each pair of regions, we obtain the eigenvalues of the estimated covariance operator of the difference of the sample means of both random samples obtained as shown in Section 3. For the Eastern-Western regions, the cumulative percentage of total variance explained by the first 10 eigenvalues are 0.8749, 0.9787, 0.9925, 0.9947, 0.9962, 0.9972, 0.9977, 0.9983, 0.9987 and 0.9989, for the Eastern-Northern regions, these are given by 0.8822, 0.9452, 0.9755, 0.9960, 0.9981, 0.9992, 0.9995, 0.9997, 0.9998 and 0.9998, while for the Western-Northern pair these are given by 0.7290, 0.9380, 0.9728, 0.9956, 0.9975, 0.9985, 0.9990, 0.9994, 0.9996 and 0.9997. As can be seen, in the three cases, the cumulative percentages grow slowly from 99%. Therefore, we select 99% of the total variation in the three cases. Table 6 shows that the value of K selected via the CPV is 3 for Eastern-Western regions and $K = 4$, otherwise. Thus, we conclude that the functional Hotelling's T^2 statistic reject the null

Figure 6: The estimated covariance operators for the three groups.

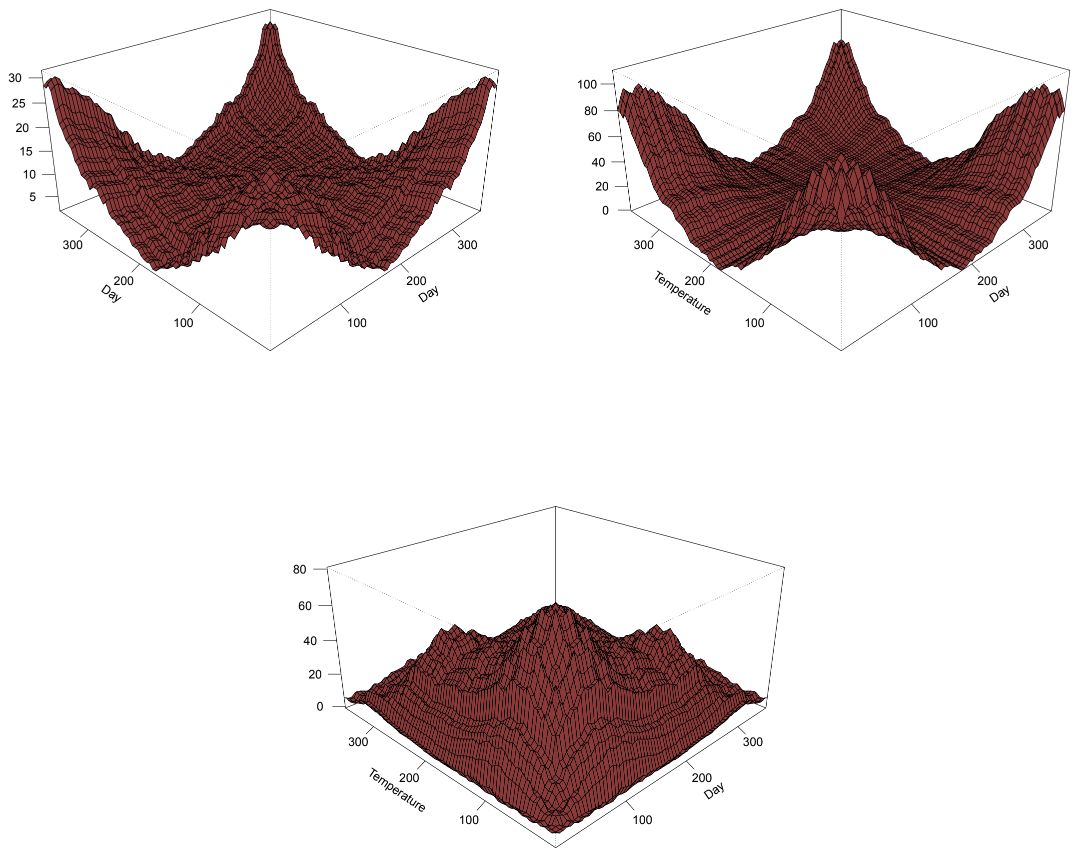


Figure 7: The contours of the estimated covariance operators for the three groups.

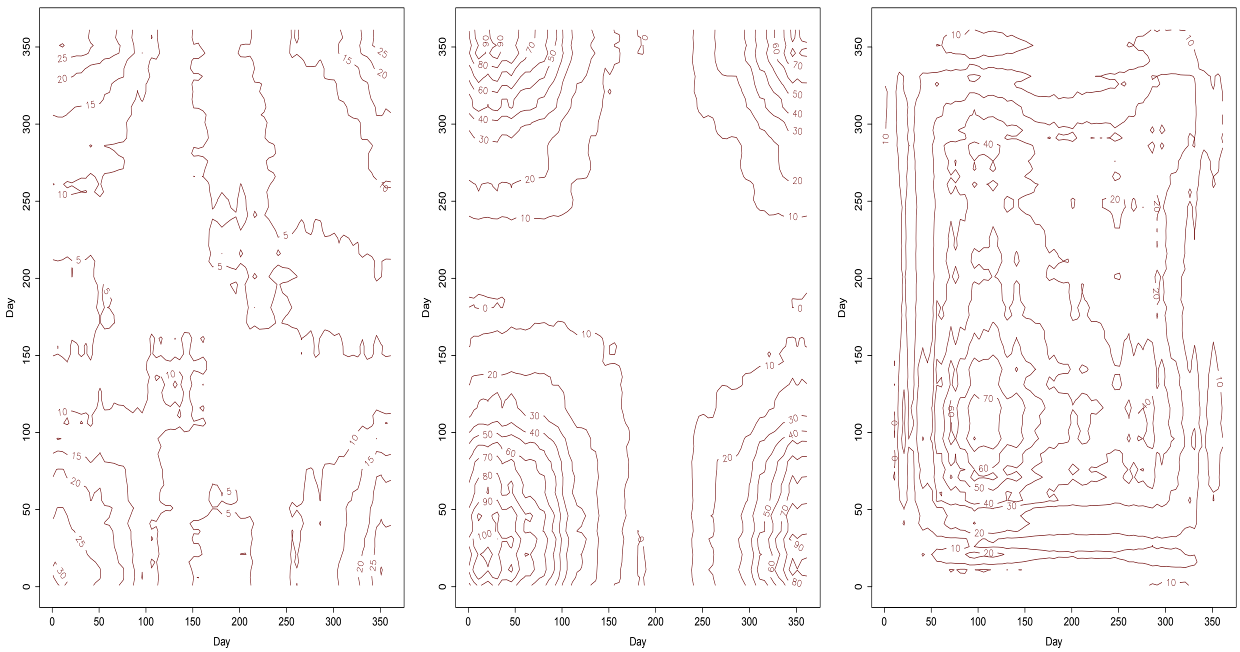


Figure 8: The first 10 eigenvalues of the estimated covariance operators for the three groups.

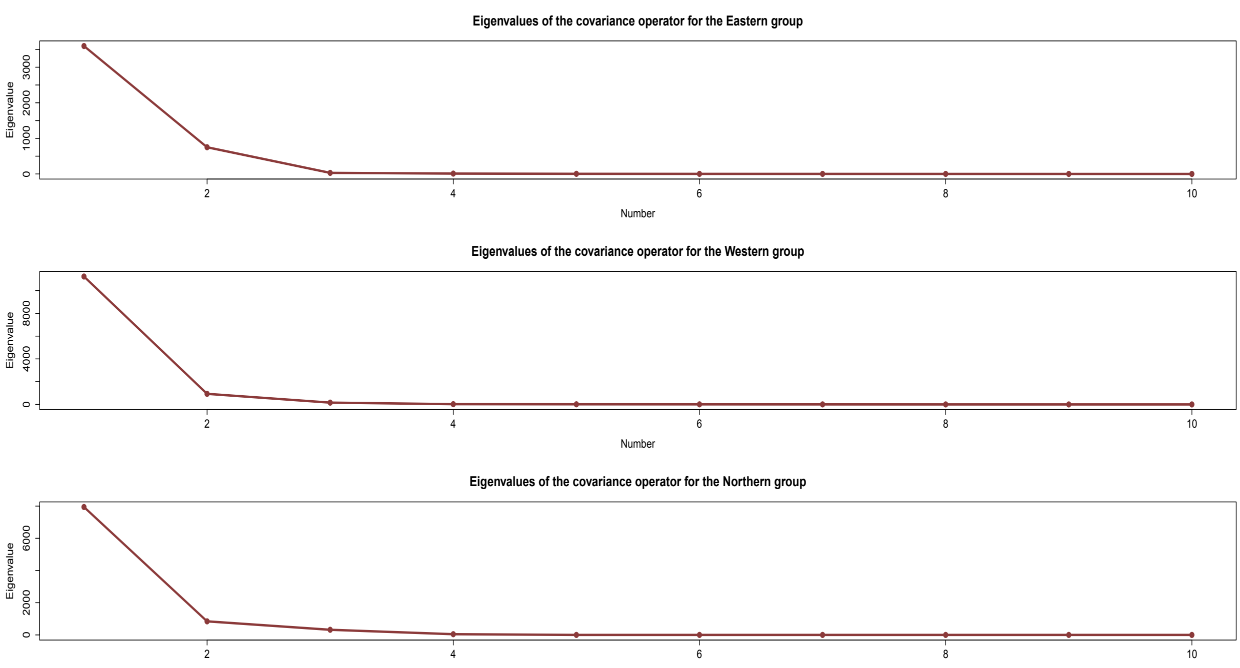


Table 6: P-values (in percent) of the tests based on statistics T^2_{FD} and U_{FD} applied to the Canadian Temperature data set for Eastern-Western, Eastern-Northern and Western-Northern stations.

	Eastern-Western		Eastern-Northern		Western-Northern	
K	T^2_{FD}	U_{FD}	T^2_{FD}	U_{FD}	T^2_{FD}	U_{FD}
1	30.97	31.13	1.17×10^{-5}	0	4.16×10^{-5}	0
2	53.73	34.17	1.10×10^{-22}	0	4.5×10^{-8}	0
3	2.58×10^{-8}	20.63	8.92×10^{-22}	0	2.17×10^{-7}	0
4	7.29×10^{-11}	19.74	1.44×10^{-26}	0	8.01×10^{-7}	0
5	9.14×10^{-15}	19.07	2.55×10^{-30}	0	1.27×10^{-9}	0
6	9.88×10^{-15}	19.52	7.80×10^{-38}	0	9.53×10^{-10}	0
7	7.19×10^{-15}	19.36	3.99×10^{-40}	0	1.00×10^{-9}	0
8	1.56×10^{-15}	19.07	9.88×10^{-139}	0	2.48×10^{-10}	0
9	1.20×10^{-26}	18.79	3.02×10^{-152}	0	1.54×10^{-10}	0
10	1.45×10^{-46}	19.17	2.68×10^{-151}	0	3.96×10^{-10}	0
11	2.32×10^{-91}	18.17	1.41×10^{-181}	0	2.07×10^{-12}	0
12	2.26×10^{-94}	17.62	1.88×10^{-246}	0	4.16×10^{-31}	0
13	9.91×10^{-100}	17.98	5.65×10^{-269}	0	9.13×10^{-36}	0
14	5.57×10^{-99}	18.01	0	0	6.04×10^{-73}	0
15	4.48×10^{-100}	17.52	0	0	1.10×10^{-82}	0
$K-CPV$	2.58×10^{-8}	20.63	1.44×10^{-26}	0	8.01×10^{-7}	0

hypothesis of equality of mean functions for Eastern-Western regions when K is properly selected.

6 Conclusions

In this paper we have derived two-sample Hotelling's T^2 statistics for testing the equality of mean functions in two samples independently drawn from two functional distributions based on the functional Mahalanobis semi-distance. In particular, in the case in which the covariance operators of the two random samples are not assumed to be the same, we have proposed a bootstrap method to estimate the covariance operator of the differences between the sample means of the two random samples. The limit distributions of the statistics under the null hypothesis are chi-squared, a result that can be established from the relationship between the proposed statistics and those based on the functional principal components semi-distance proposed in Horváth and Kokoszka (2012). Indeed, we have shown that the derived two-sample Hotelling's T^2 statistics coincide with the normalized functional principal components semi-distance statistics proposed in Horváth and Kokoszka (2012). The simulations and real data application show that the two-sample Hotelling's

T^2 statistics appears to outperform the tests based on the functional principal components semi-distance given in Horváth and Kokoszka (2012). To apply the tests, it is advisable to select the number of functional principal components used in the computations of the statistics. We propose to use the cumulative percentage of the total variance. However, other selection methods such as the Bayesian information criterion and the Akaike information criterion proposed by Li et al. (2013) could be extended to two-sample problems. This would be an objective of future work.

Appendix

Proof of (14)

Using the Fourier decomposition, the difference between the sample functional means $\widehat{\mu}_{\chi_1}$ and $\widehat{\mu}_{\chi_2}$ can be written as:

$$\widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2} = \sum_{k=1}^{\infty} \widehat{\theta}_{12k} \widehat{\psi}_k, \tag{20}$$

where $\widehat{\theta}_{12k} = \left\langle \widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2}, \widehat{\psi}_k \right\rangle$ are the scores of $\widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2}$, for $k = 1, \ldots$ Using the expressions (8) and (20), it is straightforward to show that:

$$d_{FM}^K(\widehat{\mu}_{\chi_1}, \widehat{\mu}_{\chi_2})^2 = \left\langle \widehat{\Gamma}_{K,12}^{-1/2}(\widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2}), \widehat{\Gamma}_{K,12}^{-1/2}(\widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2}) \right\rangle =$$
$$= \left\langle \sum_{k=1}^{K} \frac{1}{\widehat{\lambda}_k^{1/2}} (\widehat{\psi}_k \otimes \widehat{\psi}_k)(\widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2}), \sum_{k=1}^{K} \frac{1}{\widehat{\lambda}_k^{1/2}} (\widehat{\psi}_k \otimes \widehat{\psi}_k)(\widehat{\mu}_{\chi_1} - \widehat{\mu}_{\chi_2}) \right\rangle.$$

Now, from (7) and (20), the previous expression leads to:

$$d_{FM}^K(\widehat{\mu}_{\chi_1}, \widehat{\mu}_{\chi_2})^2 = \left\langle \sum_{k=1}^{K} \frac{1}{\widehat{\lambda}_k^{1/2}} \left[\left\langle \widehat{\psi}_k, \sum_{j=1}^{\infty} \widehat{\theta}_{12j} \widehat{\psi}_j \right\rangle \widehat{\psi}_k \right], \sum_{k=1}^{K} \frac{1}{\widehat{\lambda}_k^{1/2}} \left[\left\langle \widehat{\psi}_k, \sum_{j=1}^{\infty} \widehat{\theta}_{12j} \widehat{\psi}_j \right\rangle \widehat{\psi}_k \right] \right\rangle =$$
$$= \left\langle \sum_{k=1}^{K} \frac{\widehat{\theta}_{12k}}{\widehat{\lambda}_k^{1/2}} \widehat{\psi}_k, \sum_{k=1}^{K} \frac{\widehat{\theta}_{12k}}{\widehat{\lambda}_k^{1/2}} \widehat{\psi}_k \right\rangle = \sum_{k=1}^{K} \frac{\widehat{\theta}_{12k}^2}{\widehat{\lambda}_k}.$$

Acknowledgment

The authors acknowledge financial support from Ministerio de Ciencia e Innovación grant ECO2012-38442.

References

Benko, M., Härdle, W. and Kneip, A. (2009). Common functional principal components. Annals of Statistics, 37, 1-34.

Cuevas, A., Febrero, M. and Fraiman, R. (2004). An anova test for functional data. Computational Statistics and Data Analysis, 47, 111-122.

Fan, J. and Lin, S.-K. (1998). Test of significance when data are curves. Journal of the American Statistical Association, 93, 1007-1021.

Ferraty, F. and Vieu, P. (2006). Nonparametric functional data analysis. New York: Springer.

Galeano, P., Joseph, E. and Lillo, R. E. (2014). The Mahalanobis distance for functional data with applications to classification. Technometrics, DOI:10.1080/00401706.2014.902774.

Horváth, L. and Kokoszka, P. (2012). Inference for functional data with applications. New York: Springer.

Horváth, L., Kokoszka, P. and Reeder, R. (2013). Estimation of the mean of functional time series and a two-sample problem. Journal of the Royal Statistical Society, B, 75, 103-122.

James, G. S. (1954). Tests of linear hypotheses in univariate and multivariate analysis when the ratios of the population variances are unknown. Biometrika, 41, 19-43.

Johansen, S. (1980). The Welch-James approximation to the distribution of the residual sum of squares in a weighted linear regression. Biometrika, 67, 85-92.

Kim, S. (1992). A practical solution to the multivariate Behrens-Fisher problem. Biometrika, 79, 171-176.

Li, Y., Wang, N. and Carroll, R.J. (2013). Selecting the number of principal components in functional data. Journal of the American Statistical Association, 108, 1284-1294.

Mahalanobis, P. C. (1936). On the generalized distance in Statistics. Proceedings of the National Institute of Science, India, 2, 49-55.

Mas, A. (2007). Weak convergence in the functional autoregressive model. Journal of Multivariate Analysis, 98, 1231-1261.

Nel, D. G. and van der Merwe, C. A. (1986). A solution to the multivariate Behrens-Fisher problem. Communications in Statistics - Theory and Methods 15, 3719-3735.

Ramsay, J. O. and Silverman, B. W. (2005). Functional data analysis, Second edition. New York: Springer.

Ramsay, J. O., Hooker, G. and Graves, S. (2009). Functional data analysis with R and Matlab. New York: Springer.

Rencher, A. C. (1998). Multivariate statistical inference and applications. New York: Wiley.

Rencher, A. C. (2000). Methods of multivariate analysis, Second Edition. New York: Wiley.

Yao, Y. (1965). An approximate degrees of freedom solution to the multivariate Behrens-Fisher problem. Biometrika 52, 139-147.

Zhang, C. Q., Peng, H. and Zhang, J. T. (2010). Two samples tests for functional data. Communications in Statistics - Theory and Methods, 39, 559-578.

Zhang, J. T. and Chen, J. W. (2007). Statistical inferences for functional data. Annals of Statistics, 35, 1052-1079.

Zhang, J.T., Liang, X. and Xiao, S. (2011). On the two-sample Behrens-Fisher Problem for functional data. Journal of Statistical Theory and Practice. Volume 4, No. 4.